IJBPAS, March, 2015, 4(3): 1335-1346

ISSN: 2277–4998

International Journal of Biology, Pharmacy and Allied Sciences (IJBPAS)
'A Bridge Between Laboratory and Reader'

www.ijbpas.com

FORECAST OF HOUSEHOLD GAS SUBSCRIPTION USING A MIXTURE OF NEURAL NETWORK AND TIME SERIES

HOSSEIN OSTADI[1] AND BAHAREH JAFARI[2*]

1: Department of Economic, Dehaghan Branch, Islamic Azad University, Isfahan, Iran.

2: Department of Economic, Dehaghan Branch, Islamic Azad University, Isfahan, Iran.

***Corresponding Author: E Mail: bjafari_ele@yahoo.com**

ABSTRACT

Forecast of demand is one of the issues that successful implementation of it increases the efficiency of an organization and saves the resources. This issue is of special importance in cold regions. According to the obtained statistics in the country, there is an increasing number of people that requests to use gas and the National Gas Company is also responsible for meeting these requirements. One of the problems that the Gas Company is facing is lack of precise information of the existing demand in every cycle. It goes without saying that this will in turn cause problems such as unawareness of the number of required contractors as well as absence of a program to control the suitable reserves regarding different required gas meters and as a result unknown nature of the demands for subscription by the customers. In this study an algorithm based on a mixture of neural network models and time series in order to forecast the household gas subscription of Shahrekord for the next five years was presented. The procedure was that firstly the neural network was used in order to determine the importance of the effective factors on subscription of the household gas of Shahrekord. Six different networks were studied in this research and eventually a model with the highest reliability was designed. This model is called prone neural network and can forecast monthly subscription demand with the certainty level of 92.67%. In order to subscribe for household gas in 99% of the cases, G6 and G10 gas meters were used. The number of subscribers is specified according to the number of requests for these gas meters. The results show that among the factors taken into consideration to analyze the month, average, maximum and minimum temperatures are of the most important for G6 gas meters respectively and eventually the forecast with time series that had the least error of forecast was selected and subscription forecast for the years between 2014 and 2018 was made. The results show that the number of demands for household gas subscription in Shahrekord in 2018 will increase by 73% in comparison with 2013.

Keywords: Modelling, Neural network, time series, forecast, household gas subscription, Shahrekord

INTRODUCTION

Natural gas a clean fuel is considered as the most important source of energy with a very easy use and because of high temperature value is considered as one of the most important energy resources and one of the most important products to supply social welfare and economic development of every country. Our country, enjoying huge reserves of gas in the region and the entire world has a special privilege in comparison with the other countries. According to the obtained statistics in Iran, there are an increasing number of people that requests to use gas, and the National Gas Company is also responsible for meeting these requirements.

One of the problems that the Gas Company is facing is lack of precise information of the existing demand in every cycle. It goes without saying that this will in turn cause problems such as unawareness of the number of required contractors as well as absence of a program to control the suitable reserves regarding different required gas meters that leads to the lack of a suitable planning and the customers' dissatisfaction. On the other hand, this company has a lot of data that could be used to forecast the request for gas subscription. The main goal of a distribution system of energy is to meet the needs and satisfaction of the customer. This is of high importance in cold regions. Chahar Mahal Bakhtiari province is one of the coldest provinces of Iran that is located in a mountainous area. Due to the special geographical conditions of the province, it is highly important to supply the required energy of this province. The consumers of household gas of Chahar Mahal Bakhtiari are about 93% of the total gas consumers of the province and are considered as 52% of the total natural gas consumption of the province. Shahrekord covers 47.6% of the total household subscriptions of the province and stands the first among the consumers of the province among all cities. Shahrekord has the largest number of subscribers in the province (according to the website of statistics and reports of National Gas Company). Considering what was mentioned as above and the importance of stable supply of natural gas in this city (which is both part of the cold region and known as the biggest city of Chahar Mahal Bakhtiari province that has the largest number of the household consumers), the information about the household subscribers of Shahrekord was studied.

Research literature

Natural gas is one of the most important energy resources in the entire world and due to the increasing need of countries; this type of energy is of high importance. Thus it plays an important role with regard to the economic sectors and institutions and in order to advance the economic goals and even the political objectives of the countries. Forecast studies both at national and international levels are of much attention. Forecast of demand for gas subscription requires having a suitable model. A suitable forecast requires identifying and determining the degree of importance of the relevant factors and presentation of a suitable model that could identify these factors and determine the model of changes with the highest certainty and specify the degree of importance of the factors which is necessary. The models that could be used for this purpose can be divided into classical models (statistical methods) and non-classical or inventive models.

ARIMA time series model

ARIMA model is a mixed portable average model with a self-regressive model that is applied to the data of a time series. This model has three parts: self-regressive, compilation rank and portable average. The structure of the fundamental style of this model also consists of four stages, i.e., identification and justification of the model, estimation of the index, identification and receipt of model and approval of anticipation and logical nature of the model.

Artificial neural network model

Artificial neural networks are non-classical models. Development of computer and artificial intelligence knowledge have presented modern intelligence systems that could be used for doing some tasks that are done traditionally by statistics. As **Lee and Ou-Yang, 2009** expressed, neural network can learn from examples and find out precise functional relations among data, although the fundamental relations are unknown and describing them is difficult. Due to the fundamental difference existing between these methods and the common statistical methods, they can be classified separately. The scope of using these intelligent systems is very vast and is rapidly increasing. One of the important advantages of neural networks is that they do not need to formulize the decision-making process. That is why the neural networks can work better with the complication and uncertainty in comparison with traditional methods as these systems are designed in a way to be more similar to the human's justification performance. Use of neural networks in comparison with traditional models for backup system of decision-making will consist of saving time and costs **(Zeidan *et al*, 2011)**. The first tasks related to neural networks was done in 1947 when a neural physiologist called McCulloch and a mathematician called Walter Pitts issued their thesis under the title of how the possible functioning of neurons is. Their understanding of neural network was that connection of a set of binary decision-making units can create a network with the capacity of resolving every problem.

In 1949, Donald Hob issued a book under the title of organizing behavior and introduced education in neural networks (only from psychological point of view). **Bernard Widrow, 1959** presented two neural models under the names of ADALINE (Adaptive Linear Elements) and MADALINE (Multiple Adaptive Linear Elements). The name of these two models was taken from their use of multiple adaptive linear elements. Widrow and Heff, 1960 presented a law to update the weights (education) in neural network. The idea of this law was that when a Perceptron is activated, it has a big error and can regulate the amounts of weights in a way that error is distributed in the network or at least in neighboring Perceptrons. In the same year, a neural biologist in the name of **Franc Rosenblatt, 1962** from Cornell University started working on Perceptron. In 1974, education algorithm was introduced after issue of error by Paul Werbos. In 1984, many incidents made people interested in neural networks again. John Hopfield presented an article in national academy of science. Hopfield's goal was to present a simple model of brain, but to create a tool with higher efficiency using the bilateral lines among neurons. Before then, neurons were connected together only in one direction. At the same time Reilly and Cooper used a multilayer hybrid network where every layer had the strategy of resolving a different issue. Neural networks have witnessed much advancement since the restart of researches in early 1990s to date. Different models and educational methods were introduced and developed and neural networks were used in applications to identify models, to make functions closer and to make models for linear and non-linear dynamic systems, etc. The setting of processing units and their connections was called the network typology and has a deep effect on processing abilities of the neural network. There are generally two typologies for network that express how data should flow through input, hidden and output neurons. Used up typology and retrospective typology are called as used up network if a method could be found according to which all network knots are enumerated so that there will be no connection between knots and big figures or knots and smaller figures. All connections from one knot with a small figure are connected to knots with bigger figures. A network is called retrospective when there is not such an enumeration **(Yao, 1999)**. In used up network, the data flow is unilateral in network. One of the special structures for used up networks which is paid much attention by researchers is multilayer Perceptron structure (figure one). Usually the weight of this type of networks are trained by reducing algorithms, particularly known algorithms following distribution of error. The goal of human being or machine from the education of network to allocate correct weights to each connection is in a way that the network could make correct decisions.

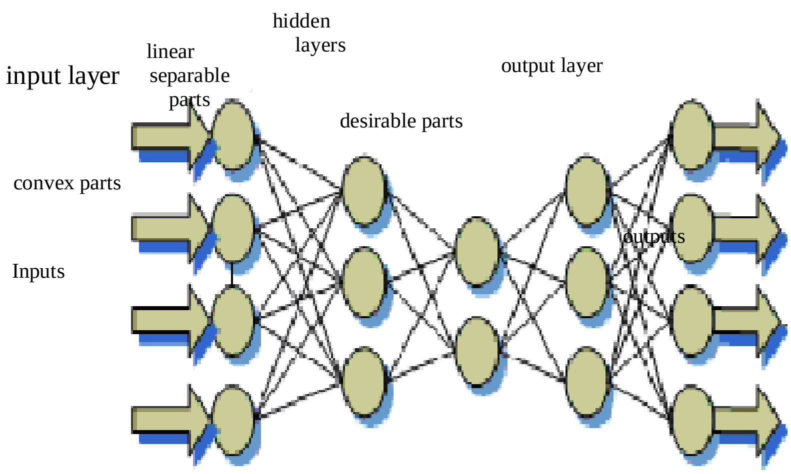


Figure 1: Multilayer Perceptron

There are usually two hidden layers which are enough to resolve any problems. However the number of more layers might increase the precision of the network in decision-making. The output layer also sends out the output amounts of the network to outside world **(Kendal *et al*, 2007)**. According to Kolmogorov Theorem, whatever is done with four layer network and done more is also achievable with a network with at most thee layers. This issue could be interpreted in a way that Perceptron is sufficient to diagnose any models in data by having at most three layers. The single layer networks can distinguish easily the linear separable models. Two layer networks can easily identify the model in convex spaces and the networks having at most three layers can easily distinguish the model in convex spaces and the networks having at most three layers can distinguish any model in any space easily **(Beale, 1998)**. Retrospective networks as their name shows are bilateral connections among processing units of networks. A subset of neurons is considered as the input processing units and special input sums are allocated to them. Then the data circulate back and forth until neurons reach a balance among the connected neighboring neurons. Retrospective networks are divided into two limited sub-groups of fully returned and returning **(Bigus, 1996)**. Used up networks have many applications in many issues nowadays and have made many achievements. The main advantage of these networks in comparison with others is the inherent generalization of it. It means that these networks can identify similar models (but not necessarily the same) from what they have learnt. Used up network models have problems that caused important discussions and challenges in this area. The most important challenges might be that there is no guarantee that the model will have a good efficiency for the existing problem or in fact shows its best efficiency. Introduction of a systematic method that could lead to the good efficiency of the network models for modeling process of networks can remove this problem to some extent. Some methods have also been presented for the issue and concentration of these methods on all aspects of network including collection of educational data, pre and post processing of data, use of different activated functions, presentation of a method for initial weighing, presentation of teaching algorithm to improve the reduction of error function in the network. What was said so far leads to efficiency of network, but the most important of them is the increasing attention to it. Network design is in other words, a suitable mixture of all the above. This issue originates from the fact that finding a suitable model will firstly increase the efficiency and secondly no suitable theoretical method has been presented in this regard **(Vosniakos, 2007)**. Designing is considered as one of the important aspects of network. In fact one of the important goals of network design methods is to find the smallest structure of network so that this structure is not only minimal, but also is able to give the fitting of a function which is given to it by the educational data. To do such a thing, the algorithms of selecting the structure should be a balance of complexity of network under construction and the amount of suitability of that network in the fitting of the function is supposed to be estimated. Design of a network can be a suitable mixture of network parameters and teaching algorithm of the network and these regulations completely depend on the issue to be resolved **(Laosiritaworn, 2009)**.

Research model

In this study, the variables related to temperature (maximum, minimum and average) and the effective temperature were extracted from the following formula:

$$T_{ef} = \left(\frac{t_{\min} + 4t_{mean} + t_{\max}}{6}\right)$$

The price of gas meter (two gas meters of G6 and G10 that 99% of the household gas meters are of these two types depending on the building flouring space), rate of construction and population as independent variables and total household gas subscribers of Shahrekord as the monthly data were defined and analyzed as dependent variables in the period from 21 March 2009 to 20 March 2014. In this study the dependent data on subscribers were extracted from the database of subscribers of Gas Company of the province and the data related to the rate of construction was taken from municipality and the temperature data was extracted from the province meteorology department. SPSS Clementine software was used for data analysis. To forecast the number of demands for subscription, the old data were used and the number of subscribers was also calculated according to the number of demands for gas meters. There are two types of gas meters in the process of subscription in household sector, i.e., G10 and G6. The summary of the neural network models situation is studied and the importance of the weight of effective variables on forecast is given in form of **Tables 1 and 2** in this method.

Table 1: Situation of different models of studied neural network for G6 gas meters

Model:	Quick Simple	Simple Dynamic	Simple multiple	Simple Prune	Simple RBFN	Simple Exhaustive prune
number of input layer of neuron	9	9	9	3	9	6
number of hidden layers	1	2	1	1	1	20
number of neurons of hidden layer 1	1:3	1:2	1:7	1:3	1:20	1:30
number of neurons of hidden layer 2	_	2:2	_	_	_	2:15
number of neurons of hidden layer 3	_	_	_	_	_	_
number of output layer of neuron	1	1	1	1	1	1
model authenticity	88.007	93.554	92.774	94.214	88.48	91.703

Table 2: Situation of different models of studied neural network for G10 gas meters

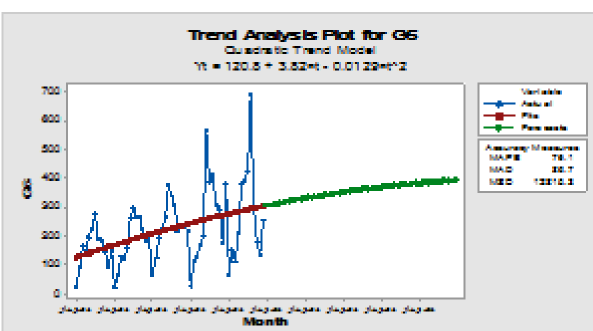
Model:	Quick Simple	Simple Dynamic	Simple multiple	Simple Prune	Simple RBFN	Simple Exhaustive prune
number of input layer of neuron	9	9	9	5	9	8
number of hidden layers	1	2	2	1	1	2
number of neurons of hidden layer 1	1:3	1:7	1:4	1:2	1:20	1:3
number of neurons of hidden layer 2	_	2:8	2:2	_	_	2:1
number of neurons of hidden layer 3	_	_	_	_	_	_
number of output layer of neuron	1	1	1	1	1	1
model authenticity	84.415	86.463	86.676	87.643	83.623	85.266

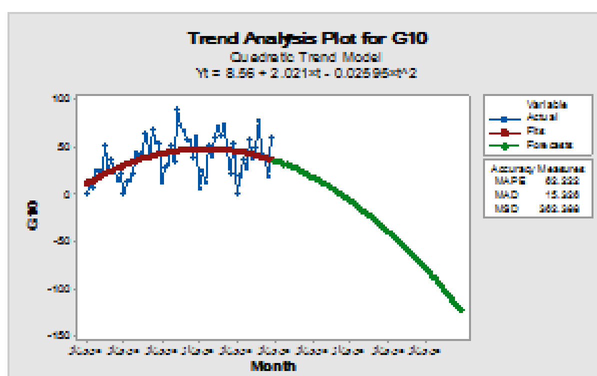
In order to forecast the gas subscription, it is necessary to seek a suitable model for each independent variable and then forecasting the independent variables, the dependent variable which is the same demand for household gas subscription and depends on the number of G6 and G10 gas meters is forecasted.

To forecast the variables related to the number of G6 and G10 gas meters by time series

As in this study, forecast of the number of household subscriptions in Shahrekord was taken into account, this forecast was made on the basis of total requested gas meters last year that was about 99% of the gas meters in household sector from G6 and G10 types. For this reason, in this study, firstly the number of these gas meters is calculated and then on the basis of this forecast the number of subscriptions in the years to come is estimated. Thus the more the precision and authenticity of anticipation regarding these two gas meters of G6 and G10 are, the more precise the forecast of the number of subscriptions in the household sector will be.

The function to forecast the number of demands for G6 and G10 gas counters with time series is as follows:





Then we will test the forecast of obtained data in neural network and will come up with the authenticity of the model as **Table 3**.

Table 3: Results to study the authenticity of forecast of G6 and G10 gas meters with time series

Model:	Simple Quick for G6	Simple Quick for G10
number of input layer of neuron	2	8
number of hidden layers	1	1
number of neurons of hidden layer 1	2	4
number of neurons of hidden layer 2	_	_
number of neurons of hidden layer 3	_	_
number of output layer of neuron	1	1
model authenticity	95.014	90.327

The first few effective parameters are air and water specifications. In future parts, the forecast of all the effective factors on the number of demands for gas meters based on time series was conducted and their chart was drawn.

Forecast of temperature: The monthly data of temperature from 1956 to 2011 (including 648 data for each specification) was studied in MINITAB software and ARIMA model was used to analyze data. To select the best ARIMA model, it is necessary to determine a few coefficients. After selecting the coefficients and comparing the parameters of errors related to remainder from forecast and reality such as SS (total square error), MS (average square error) was selected as the best model and the necessary forecasts were made on that basis. The relevant coefficients were determined according to table 3. A, B, C, D, E, F are the coefficients that are necessary to be integral numbers and can vary between one and five. To specify the model, the coefficients were shown in form of ARIMA (A,B,C) (D,E,F). Upon study of the several models, ARIMA (1, 0, 1) (5, 1, 2) model was obtained for having the highest authenticity and the least error of average temperature.

Table 4: Required data to determine an ARIMA model

	Non Seasonal	Seasonal
Auto Regrssive	A	D
Difference	B	E
Moving Average	C	F

ARIMA Model: mean mount

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	0.3514	0.1000	3.51	0.001
SAR 12	-0.6784	0.5776	-1.17	0.241
SAR 24	0.0362	0.0522	0.69	0.488

SAR 36	**0.0235**	**0.0520**	**0.45**	**0.651**
SAR 48	**-0.0674**	**0.0524**	**-1.28**	**0.199**
SAR 60	**-0.0417**	**0.0744**	**-0.56**	**0.576**
MA 1	**-0.0284**	**0.1067**	**-0.27**	**0.790**
SMA 12	**0.2559**	**0.5710**	**0.45**	**0.654**
SMA 24	**0.6859**	**0.5453**	**1.26**	**0.209**
Constant	**-0.016646**	**0.0066**	**-2.51**	**0.012**

Differencing: 0 regular, 1 seasonal of order 12
Number of observations: Original series 648, after differencing 636
Residuals: SS = 1962.16 (backforecasts excluded)
MS = 3.13 DF = 626
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	**12**	**24**	**36**	**48**
Chi-Square	**7.5**	**16.6**	**21.3**	**28.7**
DF	**2**	**14**	**26**	**38**

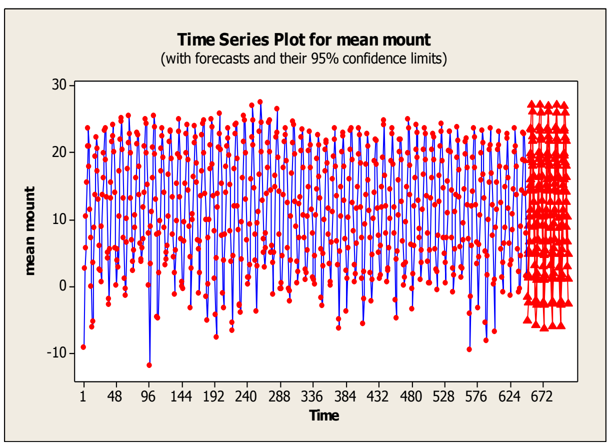


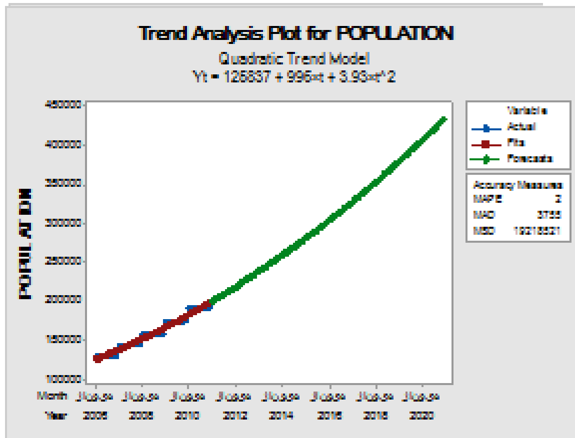
Figure 3: Chart of Meant amount of monthly temperature together with ARIMA forecast

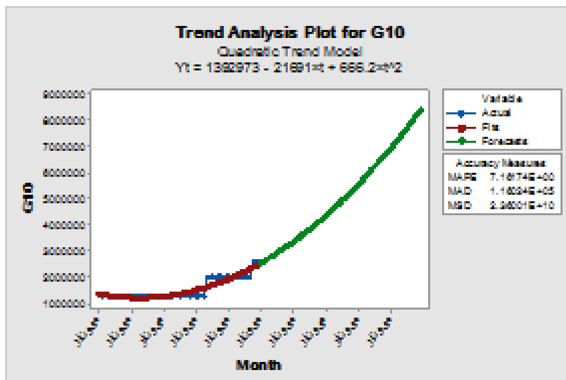
This method was also implemented for other temperature figures (effective temperature, maximum temperature and minimum temperature) and the forecasted amounts were used.

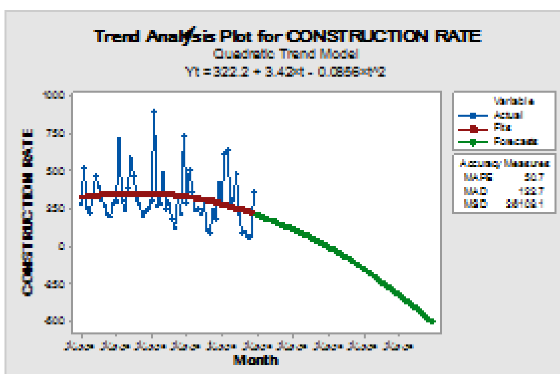
Forecast of construction rate: At this stage firstly four linear models, 2nd class, the form and function of S curve were studied. Their MSD were compared together. In the following shape, the function that has the least MSD and displays the best estimation will be shown.

Population forecast: At this stage firstly four linear models, 2nd class, the form and function of S curve were studied. Their MSD were compared together. In the following shape, the function that has the least MSD and displays the best estimation will be shown.

Display of forecast function of gas meter price with time series is as follows with the least MSD:







CONCLUSION

In this study 6 neural network models in SPSS Clementine software were studied to determine the importance of effective factors on household gas subscription of Shahrekord and eventually a model with the highest degree of certainty was selected. This model was skillful prone neural network that had one input neuron layer and one hidden layer where there were 1:2 neurons in the hidden layer and showed certainty level of 94.21% for G6 gas meter and 87.64% for G10 gas meter and the time series is more suitable and precise to estimate the number of gas meters and eventually the subscriptions.

The results show that the average, maximum and minimum temperatures, the rate of construction, the price of gas meters and population are of the most importance respectively. In order to forecast the consumed gas for each independent variable, suitable time series model was designed and then by forecast of independent variables, the dependent variable which was the actual demand for household gas subscription was forecasted. The results show that the number of demands for household gas subscription in Shahrekord will be 73% by 2018 and it will increase in comparison with 2013. What was emphasized in this study was the importance of selecting a suitable model and this selection is necessary to be based on test design and one or several criteria.

REFERENCES

[1] Kizilaslan, Recep, Karlik, Bekir, (2009). "Combination of Neural Networks Forecasters for Monthly Natural Gas", Consumption Prediction, AllBusiness.com

[2] Gorucu, F.B., Geris, P.U., Gumrah, F. (2004). Artificial Neural Network Modeling for Forecasting Gas Consumption. Energy Sources, 26(3),299-307.

[3] Gil. S. and Deferrari, J. (2004). Generalized Model of Prediction of Natural Gas Consumption. Energy Resources Technology, Vol. 126, PP. 90-98.

[4] Kendal, S.L. and Creen, M. (2007). An Introduction to Knowledge Engineering, Springer – Verlag London ltd.

[5] Vosniakos, G.C., Benardos, P.G. (2007). 'Optimizing Feed forward Artificial Neural Network Architecture', Engineering Applications of Artificial Intelligence,Vol.20, No.3, pp. 365–382.

[6] Laosiritaworn,W. and Chotchaithanakorn, N. (2009). 'Artificial Neural Networks Parameters Optimization with Design of Experiments: An Application in Ferromagnetic Materials Modeling', Chiang Mai Journal of Science, Vol. 36, No.1, pp.83-91.